Estrogen Receptor Gene Polymorphisms and Lung Adenocarcinoma Risk in Never-Smoking Women

Kuan-Yu Chen, MD, Chin-Fu Hsiao, PhD,† Gee-Chen Chang, MD, PhD,‡§ Ying-Huang Tsai, MD,‖ Wu-Chou Su, MD,¶ Yuh-Min Chen, MD, PhD,# Ming-Shyan Huang, MD, PhD,** Fang-Yu Tsai, MS,†† Shih-Sheng Jiang, PhD,†† I-Shou Chang, PhD,†††‡‡ Chih-Yi Chen, MD,§§‖‖ Chao A. Hsiung, PhD,† Chien-Jen Chen, ScD, PhD,¶¶## and Pan-Chyr Yang, MD, PhD,* and the GELAC Study Group*

Introduction: The association between estrogen receptor (ER) gene polymorphism and lung cancer risk is rarely studied. This study aimed to explore the ER gene polymorphisms associated with the lung adenocarcinoma risk in never-smoking women.

Methods: This study evaluated 532 never-smoking female patients with lung adenocarcinoma and 532 healthy controls. The *ESR1* and *ESR2* single nucleotide polymorphism (SNP) data were retrieved from a genome-wide association study. Using a multivariate-adjusted logistic regression assay, the associations of *ESR1* and *ESR2* SNPs with the lung adenocarcinoma risk were estimated. Expression quantitative trait loci analysis was performed to investigate the possible functional roles of ER gene SNPs.

Results: For *ESR1*, seven tagged SNPs were identified. Among them, rs7753153 and rs985192 were associated with lung adenocarcinoma risk (rs7753153: odds ratios [OR], 1.509; 95% confidence intervals [CI], 1.168–1.950; rs985192: OR, 1.309; 95% CI, 1.001–1.712). For *ESR2*, only rs3020450 was associated with lung adenocarcinoma risk (OR, 2.110; 95% CI, 1.007–4.422). Subjects without hormone replacement therapy (HRT) use carrying at-risk genotypes had a significantly higher lung adenocarcinoma risk than subjects with HRT carrying no at-risk genotypes (rs7753153 GG: OR, 2.133; 95% CI, 1.415–3.216; rs985192 AA/AC, OR: 1.752, 95% CI: 1.109–2.768; rs3020450 AG/GG, OR: 7.162, 95% CI: 1.608–31.90). Risk genotypes of rs7753153 ($p = 0.0248$) and rs9479122 ($p = 0.0251$) were associated with decreased *ESR1* expression.

Conclusions: ER gene SNPs are associated with lung adenocarcinoma risk in never-smoking women. The joint effects of ER gene SNPs and HRT use on lung adenocarcinoma risk highlight the importance of the gene–environment interaction in lung carcinogenesis.

Key Words: Estrogen receptor, Polymorphism, Lung cancer, Hormone replacement therapy, Carcinogenesis.

(*J Thorac Oncol.* 2015;10: 1413–1420)

Lung cancer remains the leading cause of cancer death worldwide.[1] Previous global statistics estimate that one-fourth of lung cancer patients are never smokers, including 53% of women and 15% of men.[2] Another report in Singapore demonstrated that 32.4% of non–small-cell lung cancer patients are never smokers, with 14.6% among men and 72.8% among women.[3] In that study, the major histologic type is adenocarcinoma (69.9%), especially among women (75.0%). In Taiwan, more than 90% of women lung cancer patients were never smokers, with a high percentage of adenocarcinoma (75%).[4] Lung cancer in never smokers is a unique disease entity separate from smoking-related lung cancer. Determining the risk factors associated with lung carcinogenesis in this never-smoking population has become increasingly important.

Several studies support that estrogen could promote lung cancer proliferation, either in vitro or in vivo.[5–7] Estrogens mainly interact with two subtypes of estrogen receptors (ER), ERα[8] and ERβ,[9] to exert molecular action. In breast cancer,

*Division of Pulmonary Medicine, Department of Internal Medicine, National Taiwan University Hospital and College of Medicine, Taipei, Taiwan; †Division of Biostatistics and Bioinformatics, Institute of Population Health Sciences, National Health Research Institutes, Zhunan, Taiwan; ‡Division of Chest Medicine, Department of Internal Medicine, Taichung Veterans General Hospital, Taichung, Taiwan; §Department of Medicine, School of Medicine, National Yang-Ming University, Taipei, Taiwan; ‖Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Chang Gung Memorial Hospital, Chiayi, Taiwan; ¶Division of Hematology/Oncology, Department of Internal Medicine, National Cheng Kung University, Tainan, Taiwan; #Department of Chest Medicine, Taipei Veterans General Hospital, School of Medicine, National Yang-Ming University, and Taipei Cancer Center, Taipei Medical University, Taipei, Taiwan; **Department of Internal Medicine, Kaohsiung Medical University Hospital, School of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan; ††National Institute of Cancer Research, National Health Research Institutes, Zhunan, Taiwan; ‡‡Institute of Population Health Sciences, National Health Research Institutes, Zhunan, Taiwan; §§Institute of Medicine, Chung-Shan Medical University, Taichung, Taiwan; ‖‖Department of Thoracic Surgery, Chung Shan Medical University Hospital, Taichung, Taiwan; ¶¶Genomics Research Center, Academia Sinica, Taipei, Taiwan; and ##Graduate Institute of Epidemiology, College of Public Health, National Taiwan University, Taipei, Taiwan.

Both Chen and Hsiao should be regarded as joint first authors.

Hsiung and Yang contributed equally to the work.

Disclosure: The authors declare no conflict of interest.

Address for correspondence: Pan-Chyr Yang, Department of Internal Medicine, National Taiwan University Hospital and College of Medicine, No. 7, Chung-Shan South Road, Taipei 100, Taiwan. E-mail: pcyang@ntu.edu.tw

DOI: 10.1097/JTO.0000000000000646

ISSN: 1556-0864/15/1010-1413

ERα expression is increased, and ERβ overexpression inhibits its estrodiol-stimulated cancer cell proliferation.[10] Such ERs were found in varying degrees in human non–small-cell lung cancer.[11–15] This indicated that ERα and ERβ may play a role in lung carcinogenesis.

ER gene polymorphisms have been reported to be associated with risks of breast cancer,[16,17] endometrial cancer,[18] and prostate cancer.[19] However, the association between ER gene polymorphism and lung cancer risk is seldom reported. A case–control study in the United States showed no association between *ESR2* haplotypes and lung cancer risk.[20] Another recent study in Taiwan found that ERα gene PvuII and XbaI polymorphisms are associated with non–small-cell lung cancer risk.[21]

To test the hypothesis that there may be an association between ER gene polymorphisms and lung carcinogenesis in never-smoking women, this case–control study was conducted, with *ESR1* and *ESR2* single nucleotide polymorphism (SNP) data retrieved from a genome-wide association study (GWAS) on lung adenocarcinoma in never-smoking woman.[22]

The gene–environment interaction may play a role in lung carcinogenesis. Our previous study demonstrated the impact of the interaction between hormone replacement therapy (HRT) and *epidermal growth factor receptor* (*EGFR*) polymorphism on lung adenocarcinoma risk,[23] based on the crosstalk between EGFR and ER pathways. Because estrogen may interact with ER more directly than with EGFR, we further examined potential joint effects of ER gene SNPs and HRT on the risk of lung adenocarcinoma.

METHODS

Study Population

This case–control study is a part of an ongoing cooperative study in Taiwan, the Genetic Epidemiological Study of Lung Adenocarcinoma (GELAC). From September 2002 to December 2009, patients were recruited from six medical centers. Never-smoking female patients with lung adenocarcinoma confirmed by pathologic or cytologic examination were enrolled. All subjects were Taiwanese and older than 18 years. Lung cancer histology was classified according to the World Health Organization criteria.[24] Patients with a previous history of cancer were excluded.

During the case recruitment period, control subjects were also recruited from the six medical centers. They were cancer-free individuals randomly selected from the health examination clinics of the same hospitals. Controls with a history of cancer were excluded. The control subjects were all never smokers and matched 1:1 to the case subjects based on age and sex. This study is approved by the institutional review board of each hospital.

Genotyping Analysis

The methods of genotyping were described in the previous GWAS.[24] Briefly, genomic DNA was extracted from blood samples, and the Illumina HumanHap610 Quad BeadChip on contract at deCODE Genetics in Iceland was used. A total of 550 cases and 549 controls were genotyped in the GWAS. Quality control metrics were performed to exclude SNPs if (1) there was a minor allele frequency less than 5%; (2) the call rate was less than 90%; and (3) SNPs had a missing rate between 0.02 and 0.1 and nonrandom genotype failure with p less than 0.02, and significant deviation from fitness from the Hardy–Weinberg equilibrium ($p < 0.0001$ in controls). To deal with the population substructure issue so as to avoid spurious association in conducting the GWAS,[22] we applied pairwise population concordance test in PLINK[25] to cluster individuals into homogeneous subsets and to identify outlying individuals (http://pngu.mgh.harvard.edu/purcell/plink/). Based on pairwise population concordance test, we found two clusters; a large cluster consisting of self-reported Han Chinese and a small cluster of 12 people consisting of mainly self-reported aborigines. The 12 subjects in the small cluster were excluded from further analysis. The remaining 532 cases and 532 controls with 457,504 SNPs passed quality control. The *ESR1* and *ESR2* SNPs passing the quality control were retrieved from the GWAS data. To define linkage disequilibrium patterns, SNP genotyping data from the 532 controls were uploaded to HaploView 4.2.[26]

Data Collection

The personal interview was conducted by a trained research nurse. All of the study participants provided written informed consent for blood sample collection and personal interview. Information on age; education level; body mass index (BMI, kg/m^2); active and passive cigarette smoking; cumulative duration of HRT and contraceptive medication; number of pregnancies, deliveries, and miscarriages; menopausal status; history of cooking fume exposure; history of oophorectomy; and family history of malignancies were collected.

The use of HRT was defined as a history of either estrogen replacement therapy or estrogen and progestin combination therapy for a cumulative duration of at least 3 months. Contraceptive use was defined as a history of contraceptive medication for a cumulative duration of at least 3 months. Subjects were defined as "ever cigarette smokers" if she smoked cigarettes regularly for more than 6 months.[27,28] Otherwise, they were defined as never smokers. Passive cigarette smoking was defined as inhalation of other people's cigarette smoke at the workplace or by living with family members who smoked.

The body weight of healthy controls was recorded upon enrollment. For the cases, their body weight was recorded according to the value while in a healthy state, so as to avoid underestimating the BMI because of cancer-related body weight loss. Cooking fume exposure was defined by continuous cooking for no less than 6 months. The sum of cooking fume exposure (cooking time-years) was calculated by multiplying cooking intensity (times per day) by the duration of cooking (years). A family history of breast, ovary, cervix, or endometrial cancer in first-degree female relatives was recorded.

Statistical Analysis

Differences between the case and control groups were analyzed by Pearson's χ^2 test in variables including education; passive smoking status; HRT use; contraceptive medication;

menopausal status; number of pregnancies, deliveries, and miscarriages; family history of common malignancies in women; and cooking fume exposure.

Univariate logistic regression analyses were performed to assess the effects of variables on the risk of lung adenocarcinoma. Variables with significant differences between the two groups were selected for adjustment in multivariate analysis. Odds ratios (OR) and 95% confidence intervals (CI) were calculated for logistic regression analyses.

To identify the potential SNPs associated with lung adenocarcinoma risk, a strategy based on the study by Pande et al[29] was designed with modification (Fig. 1). First, a logistic regression analyses were done to test the association between each SNP and lung adenocarcinoma, with adjustments for statistically significant covariates found by univariate analyses. Second, to filter highly correlated SNPs, those SNPs with p less than 0.05 were screened by tagging using Tagger, a program of Haploview,[26] developed by Broad Institute of MIT and Harvard (Cambridge, MA). Haplotype blocks were then constructed from the tagged SNPs. Lastly, stepwise forward selection logistic regression was performed for model selection and for selecting the tagged SNPs that had an association, while adjusting for significant covariates. The joint effects of putative high-risk SNPs genotypes in the final model and HRT use on lung adenocarcinoma risk, with adjustments for covariates, were also assessed.

The SAS system for Windows version 8.2 (SAS Institute Inc., NC) was used for all statistical analyses. All of the tests were two-sided. Statistical significance was set at p less than 0.05. For internal model validation, the bootstrap was used.[30] A bootstrap SAS Macro developed by Johnson[31] was used for calculating bootstrap estimates of OR in logistic regression. For each bootstrap estimate of OR, 10,000 bootstrap replications were conducted.

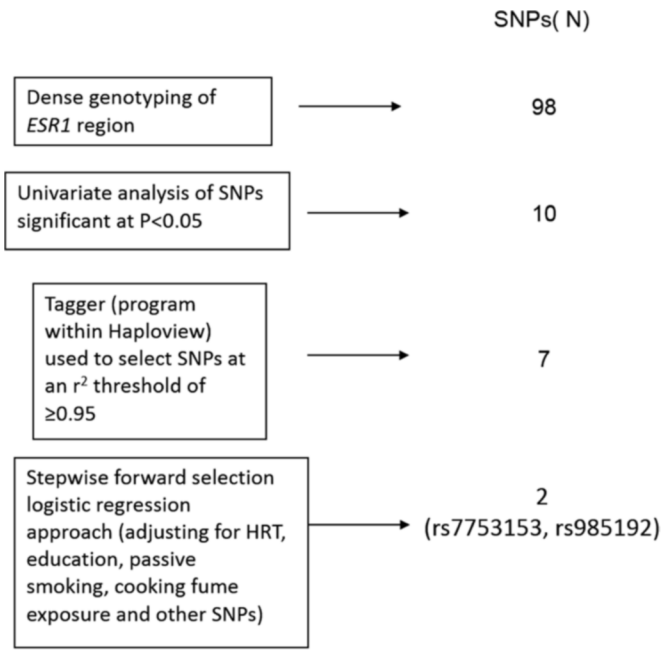

FIGURE 1. The algorithm to investigate the retrieved *ESR1* single nucleotide polymorphisms.

Expression Quantitative Trait Loci Assay

A lung cancer tissue cohort (LCTC) was complied from lung cancer patients underwent surgical operation in China Medical University Hospital, Taiwan, as part of the GELAC study supported by the National Research Program of Genomic Medicine, Taiwan. This was approved by the Institutional Review Board of the National Health Research Institutes and that of China Medical University. We obtained the informed consent of each subject in the LCTC and collected his/her tumor tissue and adjacent normal tissue, in addition to blood and clinical information.

For some of the LCTC subjects, microarray experiments were conducted to obtain their expression levels in both the tumor tissue and the adjacent normal tissue, using Illumina WG-DASL HumanRef[8] v3 or HumanHT12 v4 BeadChip. Genome-scale genotype data based on DNA from buffy coat were also obtained for some of these subjects, using Illumina Human 660W Quad BeadChip (Illumina Inc., San Diego, CA).

Preprocessing of the microarray expression data included standard quality control (per Illumina's instructions), model-based background correction,[32] and surrogate variable analysis to detect and alleviate batch effects.[33] This resulted in expression level at 24,216 probes; here, the intersection of the two Illumina probe sets was considered. Using PLINK,[25] we performed systematic quality control on the genotype data to remove poor SNPs and poor samples in the same manner as in our GWAS[22]; these resulted in 441,297 SNPs for analysis.

Finally, we obtained the genome-scale expression and genotype data of 115 nonsmoking patients with lung adenocarcinoma, referred to as the subcohort LCTCNS. The expression data used in this study can be found in the dataset Gene Expression Omnibus with accession number GSE46539. To further explore the other genes expression that may potentially associated with the risk genotype of *ESR1* and *ESR2*, expression quantitative trait loci (eQTL) analyses were performed covering the genes in 1000 kb upstream and downstream of *ESR1* and *ESR2*. Because rs9479122 is not genotyped in the patients of the LCTCNS, we obtained imputed genotype data for them, using IMPUTE2[34] and CHB[ASN] and CHS[ASN] in 1000 Genomes Project (Jun 2011 release) as reference panel; CHB[ASN] and CHS[ASN] have, respectively, 97 and 100 samples.

RESULTS

Clinical Characteristics

The demographic data of cases and control subjects were reported in a previous study concerning the interaction between *EGFR* polymorphisms and HRT use.[23] The mean age was 60.55 years for the case group, and 60.41 years for the control group ($p = 0.833$). There were significant differences the two groups in terms of the proportion of HRT use (case versus control, 25.53% versus 31.57%, $p = 0.032$), passive smoking (case versus control, 76.79% versus 70.44%, $p = 0.021$), educational level (years, >12, 7–12, <7; case versus control, 12.97% versus 17.34%; 27.44% versus 31.60%, 59.59% versus 51.06%, $p = 0.016$), and cooking fume exposure period (time-year, >110, 60–110, 10–60, ≤10; case versus control, 29.21% versus 22.68%, 26.50%

TABLE 1. The Association of Seven Tagged *ESR1* SNPs with Lung Adenocarcinoma Risk, with Bootstrap Analysis

SNP	Genotype	Case	Control	OR (95% CI)	Adjusted *P* Value[a]	Bootstrap OR[a]
rs532010	GA + GG	297 (55.83%)	331 (62.22%)	1	0.0190	1
	AA	235 (44.17%)	201 (37.78%)	1.361 (1.052–1.761)		1.363
rs9479122	AA + AG	112 (21.09%)	147 (27.68%)	1	0.0252	1
	GG	419 (78.91%)	384 (72.32%)	1.398 (1.042–1.875)		1.402
rs7753153	AA + AG	224 (42.11%)	271 (50.94%)	1	0.0008	1
	GG	308 (57.89%)	261 (49.06%)	1.544 (1.197–1.992)		1.551
rs9322336	AA	181 (34.02%)	221 (41.54%)	1	0.0139	1
	GA + GG	351 (65.98%)	311 (58.46%)	1.387 (1.069–1.799)		1.391
rs985192	CC	167 (31.39%)	195 (36.65%)	1	0.0233	1
	AA + AC	365 (68.61%)	337 (63.35%)	1.360 (1.043–1.775)		1.366
rs3003925	AA + GA	459 (86.28%)	479 (90.04%)	1	0.0418	1
	GG	73 (13.72%)	53 (9.96%)	1.511 (1.015–2.249)		1.517
rs985694	GG	140 (26.32%)	168 (31.58%)	1	0.0260	1
	AA + AG	392 (73.68%)	364 (68.42%)	1.371 (1.039–1.810)		1.379

[a]Adjusted for education level, cooking time-year, hormone replacement therapy, and passive smoking.
SNPs, single nucleotide polymorphisms; CI, confidence interval; OR, odds ratio.

versus 35.70%, 33.27% versus 28.21%, 11.03% versus 13.41%, $p = 0.002$). These variables were selected for adjustment in the multivariate analysis. No significant differences were found between the patients and controls in the other variables including BMI, menopause, contraceptives use, miscarriage, pregnancy, delivery, oophorectomy, family history of breast, cervical, ovarian, or endometrial cancers, and cooking fumes exposure.

Analysis of *ESR1* and *ESR2* SNPs Retrieved from GWAS Data

For *ESR1*, a total of 98 SNPs were retrieved and analyzed. Among them, rs532010, rs9479122, rs7753153, rs11155813, rs827423, rs9322336, rs3020317, rs985192, rs3003925, and rs985694 were associated with lung adenocarcinoma risk, without adjustment for multiple comparison effects (Supplementary Table 1, Supplementary Digital Content 1, http://links.lww.com/JTO/A864). Using the Tagger program, seven tagged SNPs were screened (Table 1), including rs532010 ($p = 0.0190$), rs9479122 ($p = 0.0252$), rs7753153 ($p = 0.0008$), rs9322336 ($p = 0.0139$), rs985192 ($p = 0.0233$), rs3003925 ($p = 0.0418$), and rs985694 ($p = 0.0260$). All bootstrap estimates of ORs were very close to estimates from the multivariate models (Table 1).

One haplotype block, block 1, was constructed to identify SNP sets in linkage disequilibrium among the seven tagged SNPs (Fig. 2). A stepwise forward selection logistic regression approach was used for the seven SNPs, with adjustment for HRT, education level, passive smoking, and cooking fume exposure. The SNPs rs7753153 and rs985192 were significantly associated with lung adenocarcinoma exposure (rs7753153: OR, 1.509; 95% CI, 1.168–1.950; rs985192: OR, 1.309; 95% CI, 1.001–1.712; Table 2). All bootstrap estimates of the ORs were very close to estimates from the multivariate models (Table 2).

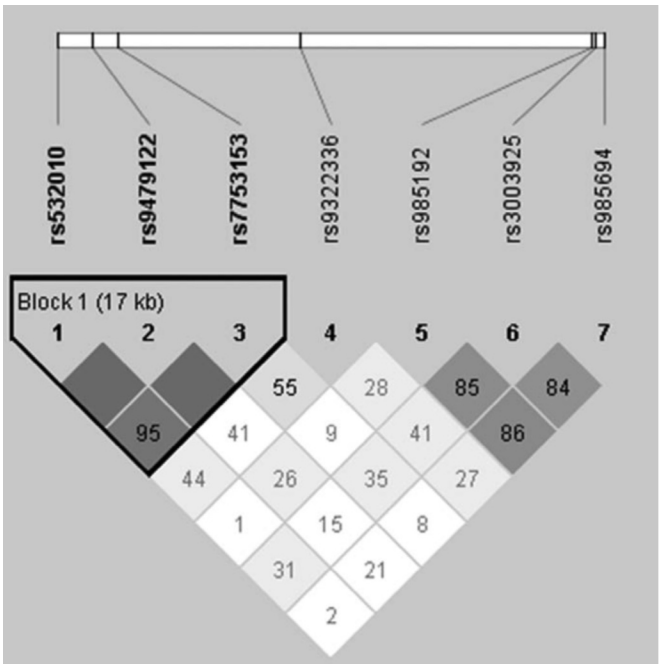

FIGURE 2. The linkage disequilibrium pattern of seven tagged *ESR1* single nucleotide polymorphisms (SNPs) interacting with hormone replacement therapy (HRT) use, denoted by r^2 values and shading. The correlation between the SNPs was indicated by the level of r^2 values and the density of shading.

For *ESR2*, 20 SNPs were retrieved and analyzed. Only rs3020450 was associated with lung adenocarcinoma risk, without adjustment for multiple comparison effects (OR, 2.110; 95% CI, 1.007–4.422; $p = 0.0478$; Supplementary Table 2, Supplementary Digital Content 2, http://links.lww.com/JTO/A865).

TABLE 2. The *ESR1* SNPs Associated with Lung Adenocarcinoma Risk Selected by a Stepwise Forward Selection Logistic Regression Approach, with Bootstrap Analysis

SNP	Genotype	OR (95% CI)	Adjusted *P* Value[a]	Bootstrap OR[a]
rs7753153	AA + AG	1	0.0017	1
	GG	1.509 (1.168–1.950)		1.513
rs985192	CC	1	0.0489	1
	AA + AC	1.309 (1.001–1.712)		1.314

[a]Adjusted for education level, cooking time-year, hormone replacement therapy, and passive smoking.

SNPs, single nucleotide polymorphisms; CI, confidence interval; OR, odds ratio.

Joint Effects of ER Gene SNP and HRT Use on Lung Adenocarcinoma Risk

The joint effects of ER SNPs and HRT use on the risk of lung adenocarcinoma are revealed in Table 3. For *ESR1*, subjects not using HRT and carrying the at-risk genotypes of rs7753153 and rs985192 had significantly higher risks of developing lung adenocarcinoma than those with HRT use and without the at-risk genotypes (no HRT + rs7753153 GG, OR, 2.133; 95% CI, 1.415–3.216; no HRT + rs985192 AA/AC, OR, 1.752; 95% CI, 1.109–2.768). In contrast, for *ESR2*, subjects without HRT use and carrying the at-risk genotypes of rs3020450 had significantly higher risks of developing lung adenocarcinoma than those with HRT use and not carrying the at-risk genotypes (no HRT + rs7753153 AG/GG, OR, 7.162; 95% CI, 1.608–31.90).

All bootstrap estimates of the ORs in *ESR*1 were very close to estimates from the multivariate models. The bootstrap estimates of OR in *ESR2* SNP were higher than the original OR, which might be due to the small sample size.

Position and Functional Location of *ESR1 and ESR2* SNPs Associated with Lung Adenocarcinoma Risk in Never-Smoking Women

The positions and functional locations of the seven tagged *ESR1* SNPs associated with lung adenocarcinoma risk are listed in Table 4. The *ESR1* SNP rs532010, rs9479122, and rs7753153, which were categorized in a haplotype block, are all located in intron 3. The SNP rs9322336 is located in intron 4, whereas rs985192, rs3003925, and rs985694 are in intron 6. The *ESR2* SNP rs3020450 is located in 5′ UTR, the promoter region.

Expression Quantitative Trait Loci Assay

We performed eQTL analysis for the seven tagged *ESR1* SNPs and the *ESR2* SNP, rs3020450. Among the *ESR1* SNPs, rs7753153 risk genotype GG and rs9479122 risk genotype GG were associated with decreased expression of *ESR1* (rs7753153, β= −0.0919, $p = 0.0248$; rs9479122, β= −0.1014, $p = 0.0251$). For *ESR2*, no association was found between the rs3020450 risk genotype and *ESR2* expression.

SNP rs7753153 risk genotype GG was associated with decreased expression of *AKAP12* (β = −0.2113, $p = 0.0110$) and increased expression of *PLEKHG1* (β=0.0641, $p = 0.0139$). SNP rs985192 risk genotype AA + AC were associated with the decreased expression of *AKAP12* (β = −0.0874, $p = 0.0308$) and increased expression of *RMND1* (β = 0.0985, $p = 0.0091$). A borderline association was found between SNPs rs985694 risk genotype AA + GA and decreased expression of *VIP* (β=−0.1520, $p = 0.0477$).

The *ESR2* SNP rs3020450 risk genotype AG + GG was associated with increased expression of *SYNE2* (β = 0.1533, $p = 0.0279$) and decreased expression of *MTHFD1*

TABLE 3. The Joint Effects of *ESR1* and *ESR2* SNP and Hormone Replacement Therapy on Lung Adenocarcinoma Risk, with Bootstrap Analysis

Gene/SNP	Genotype/HRT	Case	Control	Odds Ratio	Adjusted *P* Value[a]	Bootstrap Odds Ratio
ESR1						
rs7753153	AA/AG + HRT use	54 (10.36%)	85 (16.25%)	1		1
	AA/AG + no HRT use	166 (31.86%)	183 (34.99%)	1.439 (0.949–2.180)	0.0863	1.442
	GG + HRT use	79 (15.16%)	78 (14.91%)	1.709 (1.061–2.752)	0.0275	1.731
	GG + no HRT use	222 (42.61%)	177 (33.84%)	2.133 (1.415–3.216)	0.0003	2.154
rs985192	CC + HRT use	41 (7.87%)	56 (10.71%)	1		1
	CC + No HRT use	123 (23.61%)	136 (26.00%)	1.255 (0.770–2.045)	0.3631	1.254
	AA/AC + HRT use	92 (17.66%)	107 (20.46%)	1.269 (0.764–2.109)	0.3573	1.274
	AA/AC + No HRT use	265 (50.86%)	224 (42.83%)	1.752 (1.109–2.768)	0.0162	1.754
ESR2						
rs3020450	AA + HRT use	2 (0.38%)	14 (2.68%)	1		1
	AA + no HRT use	10 (1.92%)	9 (1.72%)	5.945 (1.028–34.38)	0.0465	26.27
	AG/GG + HRT use	131 (25.14%)	149 (28.49%)	5.789 (1.285–26.08)	0.0222	26.03
	AG/GG + no HRT use	378 (72.55%)	351 (67.11%)	7.162 (1.608–31.90)	0.0098	32.09

Subjects carrying GG genotype of rs7753153, AA/AC genotype of rs985192, and AG/GG genotype of rs3020450 were considered as at risk.

[a]Adjusted for education level, cooking time-year, and passive smoking.

SNP, single nucleotide polymorphism; HRT, hormone replacement therapy.

TABLE 4. Position and Functional Location of *ESR1* and *ESR2* SNPs Associated with Lung Adenocarcinoma Risk in Never-Smoking Women

Gene	Name	Position	Functional Location
ESR1	rs532010[a]	152172611	Intron 3
	rs9479122[a]	152182633	Intron 3
	rs7753153[a]	152189791	Intron 3
	rs9322336	152242123	Intron 4
	rs985192	152325171	Intron 6
	rs3003925	152326151	Intron 6
	rs985694	152328318	Intron 6
ESR2	rs3020450	63838055	5′ UTR

[a]Clustered in haplotype block 1.

(β = −0.0783, $p = 0.0274$), *ZBTB25* (β = −0.1169, $p = 0.0314$), and *SPTB* (β = −0.7382, $p = 0.0002$).

DISCUSSION

This study demonstrates the association between ER gene SNPs and the risk of lung adenocarcinoma in never-smoking women. In subjects without HRT use, there is a significant increase in lung adenocarcinoma risk while carrying the at-risk genotypes of ER gene SNPs.

None of the seven tagged *ESR1* SNPs has been reported to be associated with lung cancer risk. Three SNPs are clustered in a haplotype block located in intron 3, which may indicate the possible association with lung cancer risk in this region. SNP rs532010 is associated with lipolytic sensitivity to noradrenaline.[35] SNP rs9322336 is associated with increased risk of musculoskeletal toxicity-related exemestane discontinuation in breast cancer patients,[36] whereas rs985694 is associated with the risk of type two diabetes mellitus.[37] No associations were reported between these seven SNPs and breast cancer risk.

In the 20 *ESR2* SNPs retrieved, only rs3020450 is significantly associated with lung adenocarcinoma risk. This SNP is located in the promoter region of *ESR2*, which is in the 5′ region directly adjacent to the transcription start site. There are several studies on the association of rs3020450 and lipid disorders or malignancies other than lung cancer. It is associated with an increased risk of developing lipoatrophy in HIV-infected patients on Highly Active Antiretroviral Therapy.[38] However, rs3020450 has been tested in case–control studies, which showed no association with risks of endometrial carcinoma, breast cancer, and fibroid tumor.[39–42] Further studies on the influence of rs3020450 on the *ESR2* function are warranted.

In eQTL analysis for *ESR1* SNPs, the risk genotypes of rs7753153 and rs9479122 were associated with the decreased expression of *ESR1*. So far, no previous studies reported the association between rs7753153, rs9478122, and ERα expression. These risk genotypes may be associated with low ERα expression and attenuate the protective effect of HRT in never-smoking women. *AKAP12* is a gene encoding A-kinase anchoring protein 12, which is methylated in lung cancer.[43] The risk genotypes of rs7753153 and rs985192, two SNPs in the multivariate analysis model associated with lung adenocarcinoma risk, were associated with the decreased expression level of *AKAP12*. This finding may suggest the potential interaction between *AKAP12* expression and *ESR1* SNPs in lung carcinogenesis. *PLEKHG1* encodes pleckstrin homology domain-containing family G member 1. *RMND1* encodes the protein, required for meiotic nuclear division 1 homolog, which is located in the *C6ORF97-ESR1* locus associated with breast cancer susceptibility.[44] Whether *PLEKHG1* and *RMND1* were associated with *ESR1* expression is not clear.

In contrast, ESR2 SNP rs3020450 were associated with increased expression of *SYNE2* and decreased expression of *MTHFD1*, *ZBTB25*, and *SPTB*. SYNE2 encodes spectrin repeat containing nuclear envelope 2, which is a nuclear outer membrane protein binding F-actin in cytoplasm. Beta-spectrin, encoded by *SPTB*, is a cytoskeletal protein existed mainly in erythrocyte.[45] ZBTB25 is considered as a transcription factor.[46] The genetic variant of *MTHFD*1 was reported to be associated with prognosis of lung cancer patients.[47] The association between these gene expression and lung adenocarcinoma risk has not yet been studied. Further functional studies are warranted.

Recently, Schwartz et al[48] reported that HRT and estrogen use is only associated with a reduced risk of histopathologically ERα-positive and/or ERβ-positive non–small-cell lung cancer. Among postmenopausal women, none of the hormone-related variables are associated with nuclear ERα-negative and/or ERβ-negative non–small-cell lung cancer.[48] In breast cancer, a joint effect of ER sequence variants and endogenous estrogen exposure on breast cancer risk has been also reported.[49] This study demonstrates the joint effects of ER gene SNPs and HRT use on the risk of lung adenocarcinoma. When studying the association between HRT and lung cancer risk, the interaction between ER SNPs and HRT use should be considered.

In the past few years, there have been several GWASs conducted for lung cancer.[50–56] Almost all have been conducted to detect the main effect of genetic variants. Thus, few data concerning gene–environment interactions are presented in the GWASs. In our study, the cases had more exposure to fumes and passive smoking and less use of HRT. Among subjects with different environmental exposures, susceptibility alleles for lung cancer may be masked by heterogeneity. These masked genetic variants can be identified by analyzing GWAS data with environmental exposure information. Environmental exposure, which may modify cancer risk in individuals harboring certain gene polymorphisms, can be also clarified. Previously, HRT use been demonstrated to have a protective effect on lung carcinogenesis in women.[57] In this study, *ESR1* and *ESR2* SNPs have been further retrieved from GWAS data to disclose their joint effects with HRT use on lung adenocarcinoma risk. These findings may help in identifying high-risk subjects in lung cancer screening. Further studies on gene polymorphisms interact with cooking fume exposure period and passive smoking may be helpful for understanding the mechanism of lung carcinogenesis.

There are major limitations in this study. First, this study is a secondary analysis from a molecular epidemiologic study on susceptibility markers of lung cancer. The case number is relatively small in comparison with other population-based studies. However, despite its limited case number, this study is the first on the joint effects of ER gene SNPs and HRT on lung cancer risk in never-smoking women. These findings suggest that the association between ER gene SNP and lung adenocarcinoma risk may be modified by HRT use and indicate that adjustments of environmental factors, like HRT use, are important in genetic epidemiologic studies on lung carcinoma risk in women. Second, this study is lack of an independent validation from other studies. However, studies recruited never-smoking subjects in lung carcinogenesis are few. These findings should be validated by further large population-based studies on never-smoking subjects. Third, the bootstrap analysis was used for internal validation because it is often used with increasing popularity in model validation. However, the bootstrap theory involves showing consistency of the estimate, for which large sample size may be required. If the sample size is small, the bootstrap analysis may not work well because the set of possible bootstrap samples is not rich enough. It should be noted that to investigate the joint effects of SNP and HRT, the bootstrap estimates of OR for internal validation in *ESR2* SNP rs3020450 were larger than the original OR. In this case, the bootstrap estimates cannot work well because the sample size for subjects carrying AA genotype with HRT use is small.

At last, this is a hospital-based case–control study, there might be a healthy-volunteer effect for the controls recruited from the health examination department. Although one study on hormonal factors and the risk of non–small-cell lung cancer demonstrated similar results for hospital-based and population-based controls,[58] our results should be interpreted with caution. Selecting the cases only from those who previously attended the health examination departments will make cases and controls more comparable.

In conclusion, ER gene SNPs may be associated with lung adenocarcinoma risk in never-smoking women. The joint effects of ER gene SNP and HRT use highlight the importance of the gene–environment interaction in lung carcinogenesis among never-smoking women.

ACKNOWLEDGMENTS

The GELAC study was supported by the National Research Program on Genomic Medicine (NRPGM) of Taiwan under grants DOH97-TD-G-111-026 and DOH98-TD-G-111-015.

REFERENCES

1. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. *CA Cancer J Clin* 2014;64:9–29.
2. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin* 2005;55:74–108.
3. Toh CK, Gao F, Lim WT, et al. Never-smokers with lung cancer: epidemiologic evidence of a distinct disease entity. *J Clin Oncol* 2006;24:2245–2251.
4. Chen KY, Chang CH, Yu CJ, Kuo SH, Yang PC. Distribution according to histologic type and outcome by gender and age group in Taiwanese patients with lung carcinoma. *Cancer* 2005;103:2566–2574.
5. Stabile LP, Davis AL, Gubish CT, et al. Human non-small cell lung tumors and cells derived from normal lung express both estrogen receptor alpha and beta and show biological responses to estrogen. *Cancer Res* 2002;62:2141–2150.
6. Kruczynski A, Astruc J, Chazottes E, Kiss R. Characterization of the hormone-sensitivity of three human non-small-cell lung cancers grafted onto nude mice. *Oncology* 1993;50:285–292.
7. Gauchez AS, Riondel J, Fernandes-Carlos T, et al. Effect of oestrone on the natural killer (NK) cell activity, antioxidant status and tumour growth in athymic mice xenografted with human tumours. *Anticancer Res* 1996;16:853–859.
8. Green S, Walter P, Kumar V, et al. Human oestrogen receptor cDNA: sequence, expression and homology to v-erb-A. *Nature* 1986;320:134–139.
9. Kuiper GG, Enmark E, Pelto-Huikko M, Nilsson S, Gustafsson JA. Cloning of a novel receptor expressed in rat prostate and ovary. *Proc Natl Acad Sci U S A* 1996;93:5925–5930.
10. Paruthiyil S, Parmar H, Kerekatte V, Cunha GR, Firestone GL, Leitman DC. Estrogen receptor beta inhibits human breast cancer cell proliferation and tumor formation by causing a G2 cell cycle arrest. *Cancer Res* 2004;64:423–428.
11. Su JM, Hsu HK, Chang H, et al. Expression of estrogen and progesterone receptors in non-small-cell lung cancer: immunohistochemical study. *Anticancer Res* 1996;16:3803–3806.
12. Kaiser U, Hofmann J, Schilli M, et al. Steroid-hormone receptors in cell lines and tumor biopsies of human lung cancer. *Int J Cancer* 1996;67:357–364.
13. Canver CC, Memoli VA, Vanderveer PL, Dingivan CA, Mentzer RM Jr. Sex hormone receptors in non-small-cell lung cancer in human beings. *J Thorac Cardiovasc Surg* 1994;108:153–157.
14. Ollayos CW, Riordan GP, Rushin JM. Estrogen receptor detection in paraffin sections of adenocarcinoma of the colon, pancreas, and lung. *Arch Pathol Lab Med* 1994;118:630–632.
15. Cagle PT, Mody DR, Schwartz MR. Estrogen and progesterone receptors in bronchogenic carcinoma. *Cancer Res* 1990;50:6632–6635.
16. Andersen TI, Heimdal KR, Skrede M, Tveit K, Berg K, Børresen AL. Oestrogen receptor (ESR) polymorphisms and breast cancer susceptibility. *Hum Genet* 1994;94:665–670.
17. Parl FF, Cavener DR, Dupont WD. Genomic DNA analysis of the estrogen receptor gene in breast cancer. *Breast Cancer Res Treat* 1989;14:57–64.
18. Ashton KA, Proietto A, Otton G, et al. Estrogen receptor polymorphisms and the risk of endometrial cancer. *BJOG* 2009;116:1053–1061.
19. Thellenberg-Karlsson C, Lindström S, Malmer B, et al. Estrogen receptor beta polymorphism is associated with prostate cancer risk. *Clin Cancer Res* 2006;12:1936–1941.
20. Paulus JK, Zhou W, Kraft P, Johnson BE, Lin X, Christiani DC. Haplotypes of estrogen receptor-beta and risk of non-small cell lung cancer in women. *Lung Cancer* 2011;71:258–263.
21. Chang HL, Cheng YJ, Su CK, et al. Association of estrogen receptor α gene PvuII and XbaI polymorphisms with non-small cell lung cancer. *Oncol Lett* 2012;3:462–468.
22. Hsiung CA, Lan Q, Hong YC, et al. The 5p15.33 locus is associated with risk of lung adenocarcinoma in never-smoking females in Asia. *Plos Genet* 2010; 6:e1001051.
23. Chen KY, Hsiao CF, Chang GC, et al.; GELAC Study Group. EGFR polymorphisms, hormone replacement therapy and lung adenocarcinoma risk: analysis from a genome-wide association study in never-smoking women. *Carcinogenesis* 2013;34:612–619.
24. Travis WD, Brambilla E, Muller-Hermelink HK, et al. *World Health Organization Classification of Tumours. Pathology and Genetics of Tumours of the Lung, Pleura, Thymus, and Heart*. Lyon, France: IARC Press, 2004.
25. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81:559–575.
26. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21:263–265.
27. Liaw KM, Chen CJ. Mortality attributable to cigarette smoking in Taiwan: a 12-year follow-up study. *Tob Control* 1998;7:141–148.
28. Kreuzer M, Gerken M, Heinrich J, Kreienbrock L, Wichmann HE. Hormonal factors and risk of lung cancer among women? *Int J Epidemiol* 2003;32:263–271.

29. Pande M, Spitz MR, Wu X, Gorlov IP, Chen WV, Amos CI. Novel genetic variants in the chromosome 5p15.33 region associate with lung cancer risk. *Carcinogenesis* 2011;32:1493–1499.
30. Shtatland ES, Kleinman K, Cain EM. A new strategy of model building in PROC LOGISTIC with automatic variable selection, validation, shrinkage and model averaging. In SUGI '29 Proceeding, 2004: Paper 191-29. Cary, NC: SAS Institute, Inc, 2004.
31. Johnson P. StatLib, 2001. Available at ftp://sunsite.univie.ac.at/mirrors/lib.stat.cmu.edu/general/bootlogit. Accessed August 19, 2015.
32. Ding LH, Xie Y, Park S, Xiao G, Story MD. Enhanced identification and biological validation of differential gene expression via Illumina whole-genome expression arrays through the use of the model-based background correction methodology. *Nucleic Acids Res* 2008;36:e58.
33. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 2012;28:882–883.
34. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* 2009;5:e1000529.
35. Nilsson M, Dahlman I, Jiao H, Gustafsson JA, Arner P, Dahlman-Wright K. Impact of estrogen receptor gene polymorphisms and mRNA levels on obesity and lipolysis–a cohort study. *BMC Med Genet* 2007;8:73.
36. Henry NL, Skaar TC, Dantzer J, et al. Genetic associations with toxicity-related discontinuation of aromatase inhibitor therapy for breast cancer. *Breast Cancer Res Treat* 2013;138:807–816.
37. Dahlman I, Vaxillaire M, Nilsson M, et al. Estrogen receptor alpha gene variants associate with type 2 diabetes and fasting plasma glucose. *Pharmacogenet Genomics* 2008;18:967–975.
38. Gasparotto AS, Sprinz E, Lazzaretti RK, et al. Genetic polymorphisms in estrogen receptors and sexual dimorphism in fat redistribution in HIV-infected patients on HAART. *AIDS* 2012;26:19–26.
39. Lattrich C, Häring J, Schüler S, Skrzypczak M, Ortmann O, Treeck O. Polymorphisms in the promoter region of estrogen receptor β gene in endometrial cancer. *Arch Gynecol Obstet* 2014;289:631–635.
40. Chen L, Liang Y, Qiu J, et al. Significance of rs1271572 in the estrogen receptor beta gene promoter and its correlation with breast cancer in a southwestern Chinese population. *J Biomed Sci* 2013;20:32.
41. Sonestedt E, Ivarsson MI, Harlid S, et al. The protective association of high plasma enterolactone with breast cancer is reasonably robust in women with polymorphisms in the estrogen receptor alpha and beta genes. *J Nutr* 2009;139:993–1001.
42. Fischer C, Juhasz-Boess I, Lattrich C, Ortmann O, Treeck O. Estrogen receptor beta gene polymorphisms and susceptibility to uterine fibroids. *Gynecol Endocrinol* 2010;26:4–9.
43. Tessema M, Belinsky SA. Mining the epigenome for methylated genes in lung cancer. *Proc Am Thorac Soc* 2008;5:806–810.
44. Yamamoto-Ibusuki M, Yamamoto Y, Fujiwara S, et al. C6ORF97-ESR1 breast cancer susceptibility locus: influence on progression and survival in breast cancer patients. *Eur J Hum Genet* 2014;23:949–956.
45. Baines AJ. Evolution of spectrin function in cytoskeletal and membrane networks. *Biochem Soc Trans* 2009;37(Pt 4):796–803.
46. Benita Y, Cao Z, Giallourakis C, Li C, Gardet A, Xavier RJ. Gene enrichment profiles reveal T-cell development, differentiation, and lineage-specific transcription factors including ZBTB25 as a novel NF-AT repressor. *Blood* 2010;115:5376–5384.
47. Jin G, Huang J, Hu Z, et al. Genetic variants in one-carbon metabolism-related genes contribute to NSCLC prognosis in a Chinese population. *Cancer* 2010;116:5700–5709.
48. Schwartz AG, Wenzlaff AS, Prysak GM, et al. Reproductive factors, hormone use, estrogen receptor expression and risk of non small-cell lung cancer in women. *J Clin Oncol* 2007;25:5785–5792.
49. Zheng SL, Zheng W, Chang BL, et al. Joint effect of estrogen receptor beta sequence variants and endogenous estrogen exposure on breast cancer risk in Chinese women. *Cancer Res* 2003;63:7624–7629.
50. Hung RJ, McKay JD, Gaborieau V, et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature* 2008;452:633–637.
51. Amos CI, Wu X, Broderick P, et al. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat Genet* 2008;40:616–622.
52. McKay JD, Hung RJ, Gaborieau V, et al.; EPIC Study. Lung cancer susceptibility locus at 5p15.33. *Nat Genet* 2008;40:1404–1406.
53. Wang Y, Broderick P, Webb E, et al. Common 5p15.33 and 6p21.33 variants influence lung cancer risk. *Nat Genet* 2008;40:1407–1409.
54. Rafnar T, Sulem P, Stacey SN, et al. Sequence variants at the TERT-CLPTM1L locus associate with many cancer types. *Nat Genet* 2009;41:221–227.
55. Landi MT, Chatterjee N, Yu K, et al. A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated with risk for adenocarcinoma. *Am J Hum Genet* 2009;85:679–691.
56. Li Y, Sheu CC, Ye Y, et al. Genetic variants and risk of lung cancer in never smokers: a genome-wide association study. *Lancet Oncol* 2010;11:321–330.
57. Chen KY, Hsiao CF, Chang GC, et al.; GEFLAC Study Group. Hormone replacement therapy and lung cancer risk in Chinese. *Cancer* 2007;110:1768–1775.
58. Meinhold CL, de Gonzalez AB, Bowman ED, et al. Reproductive and hormone factors and the risk of nonsmall cell lung cancer. *Int J Cancer* 2011; 128:1404–1413.

